



OPEN

DATA DESCRIPTOR

DNA barcode reference library for the West Sahara-Sahel reptiles

Guillermo Velo-Antón^{1,2,3}✉, Margarida Henrique², André Vicente Liz^{2,3,4}, Fernando Martínez-Freiría^{2,3}, Juan Manuel Pleguezuelos⁵, Philippe Geniez⁶, Pierre-André Crochet⁷ & José Carlos Brito^{2,3,4}✉

DNA barcode reference libraries are now continuously produced for the tree of life, which are essential pillars for the study of biological diversity. Yet, our knowledge about global diversity is largely limited in undersampled regions such as the largest warm desert, the Sahara-Sahel. This dataset provides a DNA barcode reference library for the reptiles of the Western Sahara-Sahel (WSS) and neighbouring countries across this region. It includes 760 barcodes from 133 reptile taxa, distributed in 23 families, and covering the intraspecific diversity of some species. A total of 84 species were collected in the WSS (83% of the total reptile species richness) over 18 overland field expeditions conducted since 2003. DNA barcodes resulted in a high success rate (95%) of species identification and barcoding gap analysis highlighted the effectiveness of the COI fragment as a barcode marker for the WSS reptiles. This dataset represents a comprehensive and reliable DNA reference library for the WSS, filling an important biodiversity gap across a remote and hard-to-sample region.

Background & Summary

Global biodiversity is currently undergoing an unprecedented crisis¹ caused by the devastating effects of human activities on wildlife². Yet, the knowledge available about overall global biodiversity is very limited³ because many species have not been formally described, and certain geographic regions are still undersampled, which results in underestimation of biodiversity loss⁴. The Sahara Desert, together with the neighbouring arid Sahel, is the largest warm desert in the world and both represent two major ecoregions of the African continent, covering about 11,230,000 km²⁵. Molecular taxonomy studies developed over the last decade within the Sahara-Sahel highlighted the need of urgent research to identify the hidden vertebrate diversity present in this arid and remote region^{6,7}.

The West Sahara-Sahel (hereafter WSS) spreads across Mauritania and southern Morocco and represents a transition zone between the Palaearctic and Afro-tropic biogeographical realms⁵ (Fig. 1), acting as a biogeographic crossroad⁶. A total of 103 reptile species have been identified in the WSS (Species list⁸; IUCN Red List, The Reptile Database), with representatives from several ecoregions (e.g. Mediterranean dry woodlands, Sahara Desert, Sahelian savannahs), making it one of the richest vertebrate groups in this region. Recent molecular studies have detected cryptic diversity in several reptile groups^{9–17}, suggesting that a significant amount of diversity remains undescribed. This is mainly due to the extensive information gaps regarding local species richness and individual species across this remote and hard-to-sample area⁶, and is particularly important in the local biodiversity hotspots occurring in the mountains scattered across this region, which has been identified as priority for the conservation of Sahara-Sahel biodiversity^{18,19}.

An overall genetic assessment of the reptile diversity occurring in the WSS will help to: i) identify potential cryptic diversity; ii) study evolutionary and landscape processes associated with biodiversity distribution; and iii) contribute to the conservation planning of regional reptile diversity²⁰. For this purpose, DNA barcoding comes as a fast and cost-efficient method that uses a single, short, standardized, and highly variable

¹Universidad de Vigo, Departamento de Ecología e Biología Animal, Grupo GEA, Vigo, E-36310, Spain. ²CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, 4485-661, Vairão, Portugal. ³BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661, Vairão, Portugal. ⁴Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, 4099-002, Porto, Portugal. ⁵Departamento de Zoología, Facultad de Ciencias, Universidad de Granada, E-18071, Granada, Spain. ⁶CEFE, EPHE-PSL, CNRS, Univ Montpellier, Biogéographie et Ecologie des Vertébrés, Montpellier, France. ⁷CEFE, CNRS, Univ Montpellier, EPHE, IRD, Montpellier, France. ✉e-mail: guillermo.velo@uvigo.es; jcbrito@cibio.up.pt

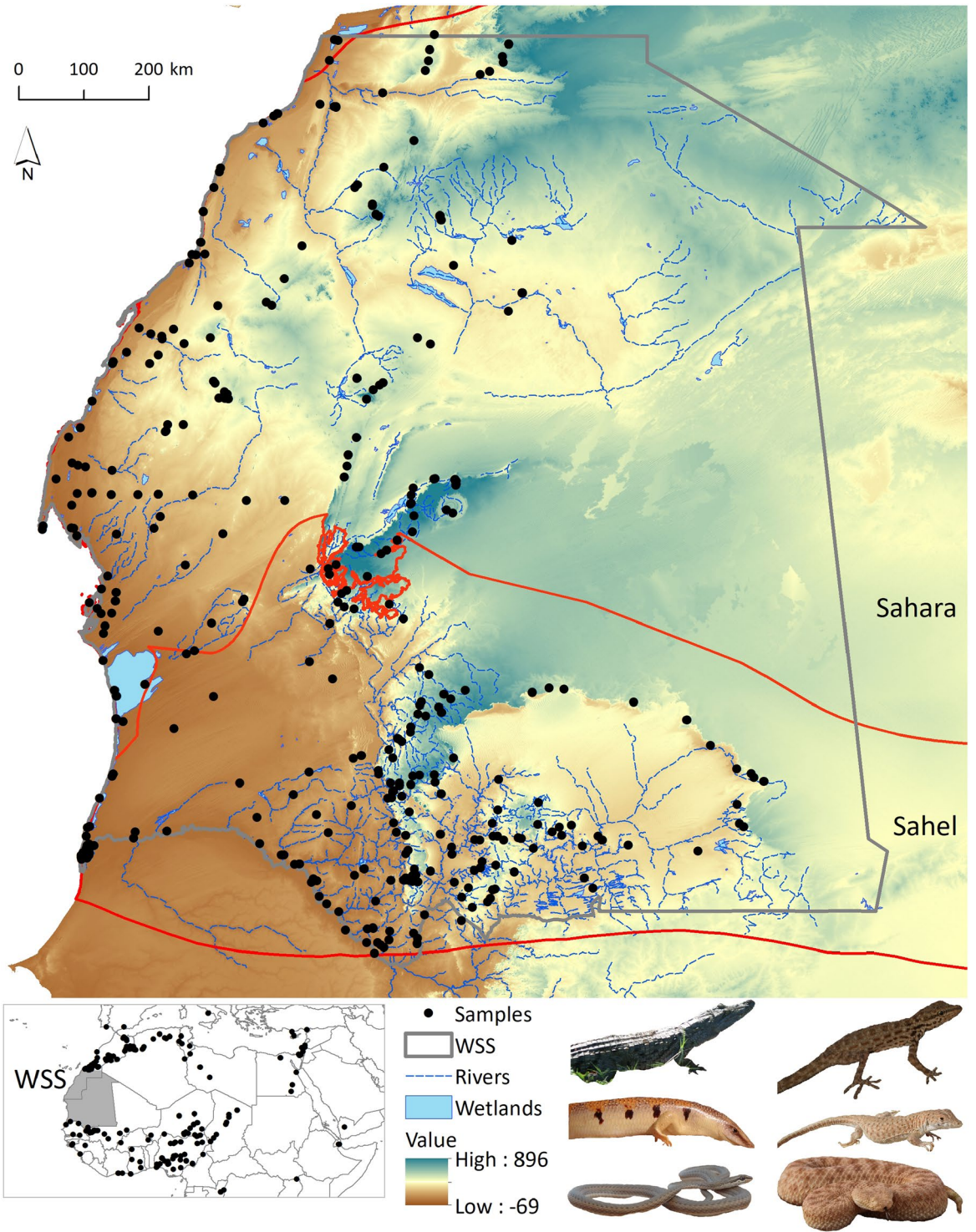


Fig. 1 Distribution of samples for the WSS reptile dataset showing the altitude across this region. The inset below indicates the distribution of samples in neighbouring countries across the Sahara and Sahel ecoregions (delimited by red lines).

genetic marker for species identification and discovery in groups where mitochondrial DNA (mtDNA) is species-specific²¹. The Consortium for the Barcode of Life (CBOL; <http://www.ibol.org/phase1/cbol/>) aims at retrieving a DNA barcode for every species on Earth and lead to the development of numerous barcoding initiatives, including the ColdCode that aims at barcoding all herpetofauna²². This initiative established the widespread use of cytochrome c oxidase 1 (COI) mitochondrial marker for barcoding studies, resulting in the discovery and assessments of herptile diversity across the globe^{20,23}, including the amphibians of the WSS⁷. DNA barcoding can thus aid to bridge the knowledge gap in poorly studied areas by providing means to understand

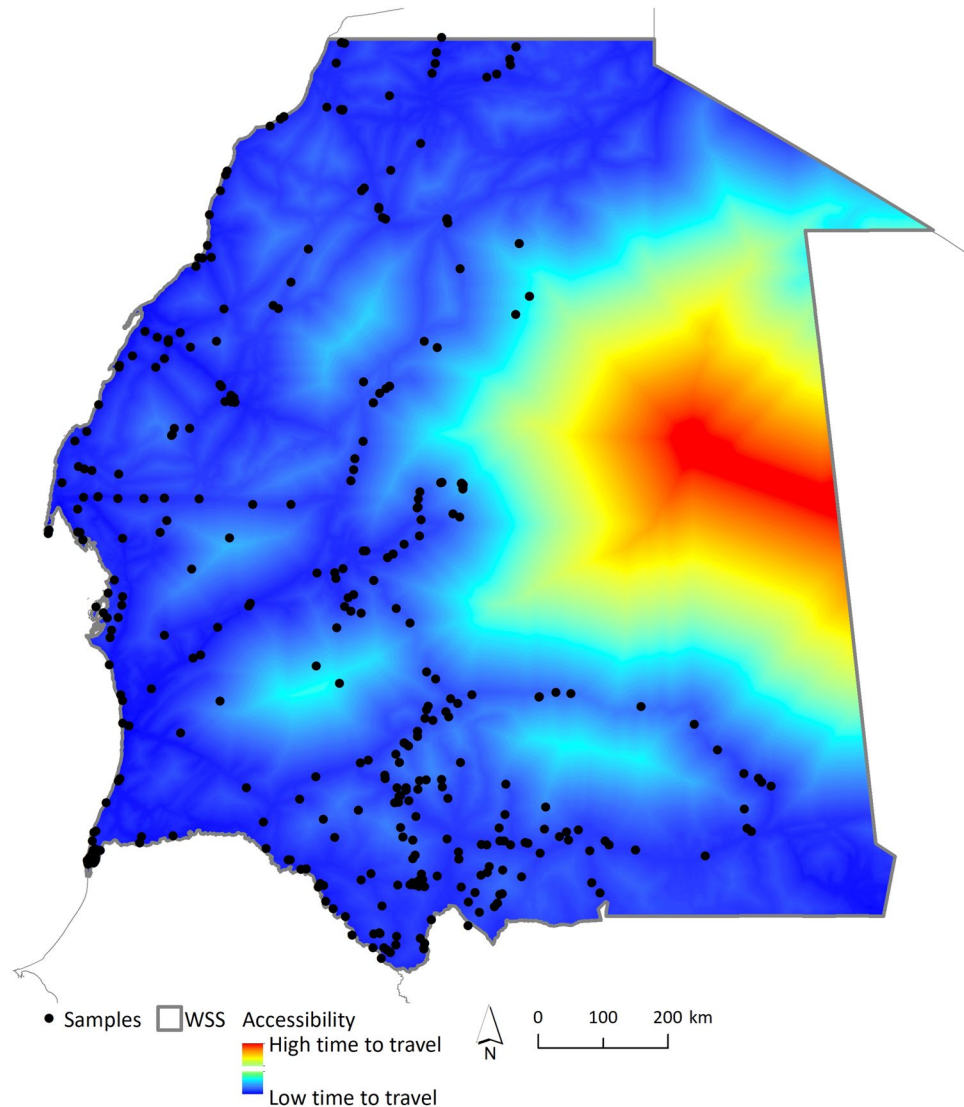


Fig. 2 Distribution of samples and accessibility to the study area. Accessibility measured as time to travel to a specific pixel (for details see Weiss *et al.*²⁸).

local species diversity²⁴. Rapid assessments of biodiversity are extremely important in current conservation decision-making²⁵.

This work represents the first DNA barcoding study of the reptiles in the WSS, which aims to: 1) establish a COI reference barcode library for the WSS reptiles; 2) assess the effectiveness of the barcode library for specimen identification using distance-based methods; and 3) identify possible candidate cryptic reptile species.

Methods

Study area. The WSS (1,024,538 km²) includes nine terrestrial ecoregions⁵ with scattered scarps-like mountains separating sandstone plateaus (Fig. 1). There is a cool, dry season from November to February and a hot, dry season from March to June. Rain falls in a single wet season from December to March in the Sahara part, and from July to September in the Sahel part, and there is a marked north-south gradient in increasing annual precipitation. Dunes, gravel and sand floodplains, compact soil, bare rock and rocky soil, grasslands, and other land-cover types cover most of WSS²⁶. The area is intrinsically remote and occasionally affected by regional conflict derived from political instability^{6,27}, which hampers regular field surveys.

Sampling strategy and collection and identification of specimens. A total of 18 overland field expeditions to WSS were conducted between 2003 and 2020 to collect samples. Field missions were developed annually, from September to December except in 2009 and 2017 (March-May) and 2015 (August). Given the remoteness of the study area and the danger of travelling in some regions (landmines from previous conflicts; Fig. 2), sampling did not follow a stratified approach but was driven by accessibility, being restricted to main routes and tracks²⁸ (Fig. 2). Visual encounter surveys in specific habitats and ad-hoc sampling (e.g. roadkill specimens) were used to find reptiles that were collected by hand or pole-noosing. For each specimen, we: 1) collected

a tissue sample from tail tip and stored it in 96% ethanol (as well as non-invasive samples, e.g. skin sheds, bones); 2) took reference digital photographs; 3) recorded the spatial location with a GPS (in WGS84 datum); and 4) preliminarily identified it to species level based on external characters following identification keys^{29–32}. After data collection, live specimens were returned to their locations. In addition, samples from WSS specimens deposited in the museum collections of MNHN Paris, BEV/CEFE Montpellier, and MHNC-UP Porto, were also included in this dataset, as well as other available samples of studied WSS taxa collected across the Sahara and Sahel ecoregions (Sample list⁸). GenBank sequences from WSS taxa, within (N = 3) or outside the study area (N = 124; 28 sequences with unknown spatial reference), and closely related species were also retrieved to aid barcoding gap analysis, which rely on the difference between maximum intraspecific and minimum interspecific genetic distances.

DNA barcode sequencing. Total genomic DNA was extracted using the QIAGEN EasySpin Kit or the QIAGEN QIAmp[®] DNA MicroKit for the samples for which the amount of tissue was limited. For samples where we expected DNA of low quality (tissue collected from dead animals or shed skin), the QIAGEN DNeasy Blood & Tissue Kit was used following manufacturer's instructions. DNA from museum samples was extracted following an optimized protocol³³. DNA extractions and subsequent procedures (PCR) were performed in sterile and isolated rooms under special conditions optimized for the manipulation of low-quality DNA.

We used the cytochrome c oxidase 1 (COI) mitochondrial marker, which is the standard barcoding marker for animals (BOLD, <http://www.boldsystems.org>). A COI fragment was amplified using degenerate primers RepCOI-F (primer forward, 5'-TNTTMTCAACNAACCACAAAGA-3') and RepCOI-R (primer reverse, 5' ACTTCTGGRTGKCCAAARAATCA-3')³⁴, except for samples of *Acanthodactylus scutellatus*, *Pristurus adrarensis* and *Philochortus zolii*, and some samples of *Acanthodactylus longipes*, *Acanthodactylus taghitensis* and *Chalcides delislei* where COI was amplified with the universal primers LCO1490 (primer forward, 5'-GGTCAACAATCATAAAGATATTGG-3') and HC02198 (primer reverse, 5'-TAAACTTCAGGGTGACCAAAAAATCA-3')³⁵. A touchdown PCR was performed with the following conditions: initial denaturation at 95 °C for 10 min, followed by an initial phase of 9 cycles of 40 s of denaturation at 95 °C, 30 s of annealing at 52 °C with a decrease in the annealing temperature by 0.5 °C per cycle until the 48 °C, and extension at 72 °C for 45 s, and then followed by a second phase with 31 cycles of 40 s of denaturation at 95 °C, 30 s of annealing at 48 °C, and elongation during 45 s at 72 °C, and a final extension cycle of 10 min at 72 °C. Adjustments of the temperature gradients were done for some species. Quality of PCR products were checked by visual examination in electrophoresis using 2% agarose gel. PCR products were outsourced for Sanger sequencing to Beckman Coulter Genomics (Essex, UK). The sequence chromatograms were visually inspected, assembled, and edited using Geneious Pro v.4.8.5 (Biomatters Ltd.). Sequences were aligned using the MUSCLE version implemented in Geneious Pro v.4.8.5 (Biomatters Ltd.) under default settings³⁶. All sequences were translated into amino acids to aid the alignment, and were checked for stop codons to detect the presence of nuclear DNA pseudogenes (NUMTs). Once the absence of NUMTs were confirmed, the sequences were trimmed to the same length.

Data Records

This reference library contains the following information: 1) Specimen ID; 2) Species ID; 3) Georeferenced data (latitude and longitude in decimal degrees) associated to each specimen collected in the field; 4) sampling date; 5) DNA barcode of each specimen; 6) Institution storing vouchers or tissue/DNA for each specimen; and 7) Photographs of live specimens collected in the field. A specimen was considered as reference for subsequent analysis and interpretation of results when the same identification was obtained from both molecular and morphological assessments. It also contains notes on the taxonomic ID for those specimens in which molecular and morphological identifications disagreed. All data associated with this study is hosted at Figshare⁸. DNA barcodes are available in GenBank (accession numbers ON943478-ON944026) and in BOLD (Ref: REWSS).

A total of 760 barcode sequences from 133 reptile taxa distributed in 23 families were analysed (Sample list⁸). These included: 1) 419 samples from 63 species known to occur in the WSS, collected within the study area; 2) 285 samples of WSS taxa collected outside the study area or lacking spatial reference; and 3) 56 samples from 46 outgroups (i.e. sister/close neighbours of WSS taxa). Samples were selected for analyses in order to cover the described taxonomic diversity and known geographic distribution of each taxon within the WSS, and represent 83% of the total reptile species richness described in the region^{29–32} (Sample list⁸). A total of 472 specimens were identified morphologically to the species level based on external diagnostic characters, from which 376 specimens were sampled in the WSS (Sample list⁸). Specimens where identification to species was not possible (documentation too poor or species pairs too difficult to identify from photos) were not treated as reference samples.

Technical Validation

We conducted independent morphological identifications by two of the authors (P.A. Crochet and P. Geniez) without regarding the genetic identification nor the geographic origin of the sample. Then, after listing the disagreements in identifications, the sequences and the voucher specimens or pictures were revised a second time to search for possible mistakes in the original identifications. Obvious mistakes in identification or in curation (mixing of photos for example) were corrected, in all other cases the mismatch between genetic and morphological identification was recorded as such.

To assess the robustness of our library and detect potential cryptic diversity in our dataset (excluding outgroups), we first evaluated the existence of a barcoding gap. A barcoding gap exists when the maximum intraspecific distance of each species is lower than its minimum distance to the nearest neighbour, and thus allows to evaluate the performance of the COI marker as a barcode, but also unveil cryptic diversity. A pairwise distance matrix was first calculated between sequences using the Kimura 2-parameter model (K2P)³⁷ to

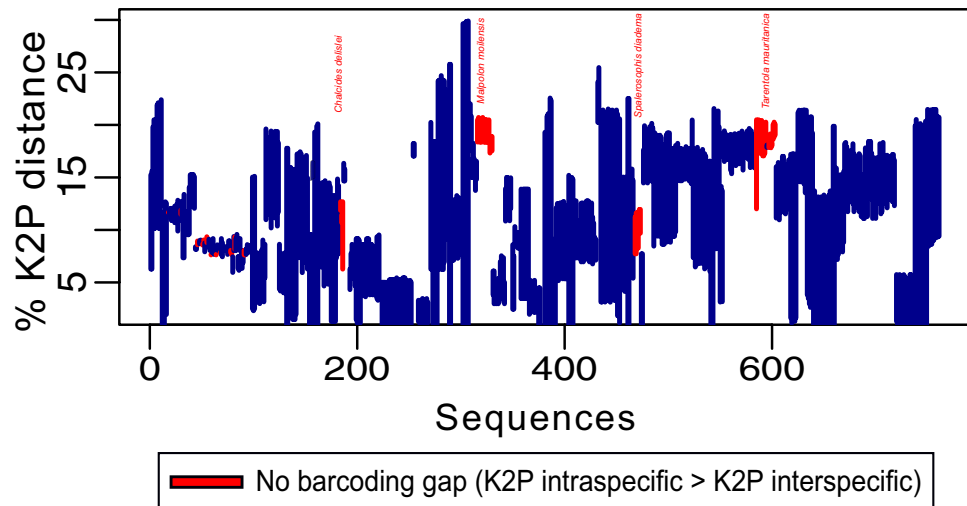


Fig. 3 Representation of the barcoding gap for the WSS reptile dataset. Each individual in the dataset is represented by a vertical line, with the top of the line representing the minimum interspecific distance and the bottom of the line representing the maximum intraspecific distance. Barcoding gaps are present if the maximum intraspecific distance is lower than the minimum interspecific distance. Individual lines are color-coded based on the absence (red) or presence (blue) of a barcoding gap. Species for which there is no barcoding gap in all sampled individuals are denoted in red.

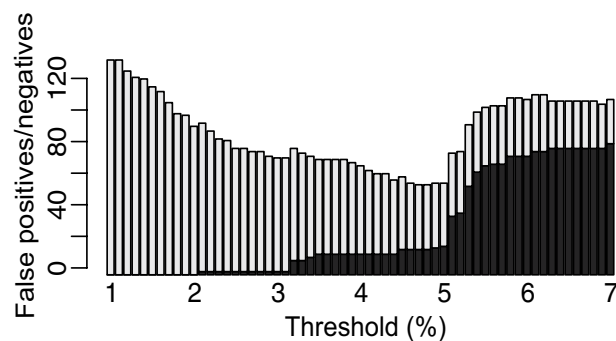


Fig. 4 Histogram illustrating the false positive (light grey) and false negative (dark grey) rate of identification of reptiles as pre-set thresholds change.

estimate the largest intraspecific distance and the smallest interspecific distance using the statistics *maxInDist* and *nonConDist*, respectively, which are implemented in the R package *spider* v.1.3³⁸. We found the presence of a barcoding gap in 92% of the samples, with a lack of barcoding gap in *Chalchides delislei*, *Malpolon moilensis*, *Spalerosophis diadema* and *Tarentola mauritanica*. We also identified a lack of barcoding gap in samples of *Acanthodactylus longipes*, *Agama agama*, *Dasypeltis scabra* and *Spalerosophis diadema* (Fig. 3; Barcoding gap⁸). Intraspecific genetic distances ranged from 0% to 20.7% (*Malpolon moilensis* showing the highest intraspecific distance). A 2.1% minimum interspecific divergence was found between *Crocodylus suchus* and *C. niloticus*.

We also tested for barcoding efficiency (i.e. assignment of barcodes to the species level) in our dataset (excluding outgroups) using two query identification analyses based on genetic distance thresholds. We used the BOLD and Meier's best close match functions (Meier's BCM³⁹) as implemented in *spider*. These two methods use a threshold based criterion that compares all specimens within the threshold of the query, and then assigns a diagnosis to each identification query: i) "correct match" (within the threshold of the query all matches are the same species); ii) "incorrect match" (i.e. closest match is a different species of the query); iii) "ambiguous match" (i.e. both correct and incorrect species matches within the threshold), and (iv) "no identification" (i.e. no species is identified within the given threshold). We explored a range of threshold values (1%–7%) before choosing the threshold value that minimized the cumulative error (false negative + false positive). Preliminary analyses indicated 5% as the most suited threshold for specimen identification (Fig. 4), which we applied to both methods. We removed species ($N = 47$) represented by only one sample (singletons) from these analyses and the outgroup samples for this count. BOLD identified 393 (94.5%), 10 (2.4%) and 13 (3.1%) samples from WSS taxa as correct, ambiguous or not identified, respectively. Meier's BCM identified 414 (99.5%) and 2 (0.5%) WSS samples as correct and incorrect, respectively (Fig. 5). Samples identified as incorrect correspond to single samples of

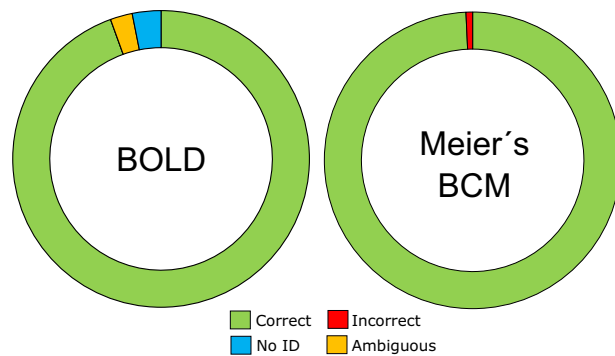


Fig. 5 Results from barcoding efficiency methods (BOLD and Meier's BCM) to determine the consistency of DNA barcodes with currently accepted taxonomy for the WSS reptile dataset.

Chalcides delislei and *Tarentola mauritanica* from WSS, and *Spalerosophis diadema* and *Dasypletilis scabra* from outside WSS (BOLD and Meier's BCM⁸).

Usage Notes

We provide a comprehensive and publicly available DNA barcode library for the West Sahara-Sahel reptile taxa that will allow barcoding or metabarcoding surveys for specimen identification, as well as for biogeographic and evolutionary studies encompassing this region.

Overall, this work improved the current knowledge on species presence, range distribution and levels of genetic structure for WSS reptile fauna. Several results can be highlighted:

A barcoding gap was largely present in our dataset, highlighting the effectiveness of the COI fragment as a barcode marker.

The lack of barcoding gap in a handful of species and the high levels of intraspecific diversity unveiled in a high number of taxa pinpoints the urgent need of further studies and taxonomic re-evaluation of some groups across the WSS, as it has been recently done for other reptiles (e.g. *Mesalina*⁴⁰). Our results suggest potential cryptic diversity at least in the following taxa (showing > 10% of intraspecific divergence): *Acanthodactylus boskianus*, *Acanthodactylus longipes*, *Acanthodactylus taghitensis*, *Agama boulengeri*, *Chalcides delislei*, *Chalcides ocellatus*, *Latastia longicaudata*, *Lytorhynchus diadema*, *Malpolon moilensis*, *Mesalina guttulata*, *Ptyodactylus oudrii*, *Spalerosophis diadema*, *Stenodactylus mauritanicus*, *Stenodactylus sthenodactylus*, *Tarentola chazaliae*, *Tarentola ehippiata*, *Tarentola mauritanica*, *Tarentola parvicarinata*, *Trachylepis perrotetii* and *Tropicolotes tripolitanus*.

This COI database also contains own and retrieved sequences from GenBank from WSS reptile taxa outside the study region, as well as phylogenetically close neighbours of WSS reptiles. Thus, this reference library is also expected to benefit a large community of researchers studying reptiles across remote and hard-to-sample areas in the vast Sahara-Sahel region.

The main issue with mtDNA barcoding as a tool for specimen identification is the possibility that mtDNA is not species-specific, due to either lack of complete lineage sorting or mtDNA gene flow (introgression) after speciation. Very few barcoding studies explicitly examine this issue by careful identification of specimens independently of genetic results. Here, we evaluated the diagnostic power of COI barcodes by comparing morphological and barcode identification in most specimens (labelled "reference" in Sample list⁸). We confirmed that COI is a reliable tool for specimen identification in most species of reptiles in our dataset.

In a few species pairs (e.g. *Stenodactylus sthenodactylus* and *S. mauritanicus*) or species complexes (e.g. *Acanthodactylus scutellatus* complex), reliable morphological identification from photographs is not possible in all specimens and discordance are more likely to result from morphological misidentifications than mtDNA lineage sharing. Indeed, a recent study combining mtDNA and nuclear data found no instance of mtDNA lineage sharing between species in the *A. scutellatus* complex.

A few (probably) real discordances between morphological and molecular identification remain after discarding possible morphological misidentifications. In one case, such discordance suggests mitochondrial introgression events between sister taxa across contact zones: *Uromastyx nigriventris* and *U. dispar* were found to be non-monophyletic in COI barcodes and although morphological identification of these two species is challenging, the sample suggesting mismatch had a seemingly typical morphology and comes from the contact zone between the two species so the lineage sharing is probably genuine (but would need to be confirmed with nuclear markers). In other cases, the discordances involve taxa where species-level systematics and species limits remain poorly understood and our results are difficult to interpret: *Trapelus boehmei* and the extralimital *T. mutabilis* or the *Tarentola ehippiata* complex (where the subspecies *hoggarensis* is most likely a valid biological species).

Last, nomenclatural instability is another source of confusion when using barcodes for specimen identification. Our samples of *Agama agama* would match samples of *Agama picticauda* in Genbank because of the confused nomenclatural situation in the *Agama agama* complex. We totally agree with the interpretations of a molecular study⁴¹ that considers the neotype designation for *Agama agama*⁴² as invalid, and accept their

conclusion on the origin of the lectotype of *A. agama*. As a consequence, we regard *Agama agama* as the valid name for the widespread species of the complex in Western African and treat *Agama picticauda* as a junior synonym of *Agama agama*.

Code availability

All analyses are implemented using the code available in the R package *spider* v.1.3 (Brown *et al.*³⁸).

Received: 6 May 2022; Accepted: 19 July 2022;

Published online: 01 August 2022

References

- Ceballos, G. *et al.* Accelerated modern human – induced species losses: entering the sixth mass extinction. *Sci. Adv.* **1**, e1400253, <https://doi.org/10.1126/sciadv.1400253> (2015).
- Hoffmann, M. *et al.* The impact of conservation on the status of the world's vertebrates. *Science* **330**, 1503–1509, <https://doi.org/10.1126/science.1194442> (2010).
- Whittaker, R. J. *et al.* Conservation biogeography: assessment and prospect. *Divers. Distrib.* **11**, 3–23, <https://doi.org/10.1111/j.1366-9516.2005.00143.x> (2005).
- Dirzo, R. & Raven, P. H. Global state of biodiversity and loss. *Annu. Rev. Env. Resour.* **28**, 137–167, <https://doi.org/10.1146/annurev.energy.28.050302.105532> (2003).
- Dinerstein, E. *et al.* An ecoregion-based approach to protecting half the terrestrial realm. *BioScience* **67**, 534–545, <https://doi.org/10.1093/biosci/bix014> (2017).
- Brito, J. C. *et al.* Unravelling biodiversity, evolution and threats to conservation in the Sahara-Sahel. *Biol. Rev.* **89**, 215–231, <https://doi.org/10.1111/brev.12049> (2014).
- Sampaio, M. *et al.* Beyond the comfort zone: amphibian diversity and distribution in the West Sahara-Sahel using mtDNA and nuDNA barcoding and spatial modelling. *Conserv. Genet.* **22**(2), 233–248, <https://doi.org/10.1007/s10592-021-01331-8> (2021).
- Velo-Antón, G. *et al.* DNA barcode reference library for the West Sahara-Sahel reptiles, *figshare*, <https://doi.org/10.6084/m9.figshare.20338335> (2022).
- Carranza, S., Arnold, E. N., Geniez, P., Roca, J. & Mateo, J. A. Radiation, multiple dispersal and parallelism in the skinks, *Chalcides* and *Sphenops* (Squamata: Scincidae), with comments on *Scincus* and *Scincopus* and the age of the Sahara Desert. *Mol. Phyl. Evol.* **46**, 1071–1094, <https://doi.org/10.1016/j.ympev.2007.11.018> (2008).
- Gonçalves, D. V. *et al.* Phylogeny of North African *Agama* lizards (Reptilia: Agamidae) and the role of the Sahara desert in vertebrate speciation. *Mol. Phyl. Evol.* **64**, 582–591, <https://doi.org/10.1016/j.ympev.2012.05.007> (2012).
- Gonçalves, D. V. *et al.* The role of climatic cycles and trans-Saharan migration corridors in species diversification: biogeography of *Psammophis schokari* group in North Africa. *Mol. Phyl. Evol.* **118**, 64–74, <https://doi.org/10.1016/j.ympev.2017.09.009> (2018).
- Gonçalves, D. V. *et al.* Assessing the role of aridity-induced vicariance and ecological divergence in species diversification in North-West Africa using *Agama* lizards. *Biol. J. Linn. Soc.* **124**, 363–380, <https://doi.org/10.1093/biolinnean/bly055> (2018).
- Metallinou, M. *et al.* Conquering the Sahara and Arabian deserts: Systematics and biogeography of *Stenodactylus* geckos (Reptilia: Gekkonidae). *BMC Evol. Biol.* **12**, 258, <https://doi.org/10.1186/1471-2148-12-258> (2012).
- Metallinou, M. *et al.* Species on the rocks: Systematics and biogeography of the rock-dwelling *Ptyodactylus* geckos (Squamata: Phyllodactylidae) in North Africa and Arabia. *Mol. Phyl. Evol.* **85**, 208–220, <https://doi.org/10.1016/j.ympev.2015.02.010> (2015).
- Kapli, P. *et al.* Historical biogeography of the lacertid lizard *Mesalina* in North Africa and the Middle East. *J. Biogeog.* **42**, 267–279, <https://doi.org/10.1111/jbi.12420> (2015).
- Tamar, K., Geniez, P., Brito, J. C. & Crochet, P. A. Systematic revision of *Acanthodactylus busacki* (Squamata: Lacertidae) with a description of a new species from Morocco. *Zootaxa* **4276**(3), 357–386, <https://doi.org/10.11646/ZOOTAXA.4276.3.3> (2017).
- Velo-Antón, G., Martínez-Freiria, F., Pereira, P., Crochet, P.-A. & Brito, J. C. Living on the edge: ecological and genetic connectivity of the Spiny-footed lizard, *Acanthodactylus aureus*, confirms the Atlantic Sahara desert as biogeographic corridor and centre of lineage diversification. *J. Biogeog.* **45**, 1031–1042, <https://doi.org/10.1111/jbi.13176> (2018).
- Vale, C. G., Pimm, S. L. & Brito, J. C. Overlooked mountain rock pools in deserts are critical local hotspots of biodiversity. *PLoS ONE* **10**, e0118367, <https://doi.org/10.1371/journal.pone.0118367> (2015).
- Brito, J. C. *et al.* Conservation Biogeography of the Sahara-Sahel: additional protected areas are needed to secure unique biodiversity. *Divers. Distrib.* **22**, 371–384, <https://doi.org/10.1111/ddi.12416> (2016).
- Hawltischek, O. *et al.* Comprehensive DNA barcoding of the herpetofauna of Germany. *Mol. Ecol. Res.* **16**, 242–253, <https://doi.org/10.1111/1755-0998.12416> (2016).
- Hebert, P. D. N., Cywinska, A., Ball, S. L. & Jeremy, R. Biological identifications through DNA barcodes. *P. Roy. Soc. Lond. B Bio.* **270**, 313–321, <https://doi.org/10.1098/rspb.2002.2218> (2003).
- Murphy, R. W. *et al.* Cold Code: the global initiative to DNA barcode amphibians and nonavian reptiles. *Mol. Ecol. Res.* **13**, 161–167, <https://doi.org/10.1111/1755-0998.12050> (2013).
- Vasconcelos, R. *et al.* Unexpectedly high levels of cryptic diversity uncovered by a complete DNA barcoding of reptiles of the Socotra Archipelago. *PLoS ONE* **11**, e0149985, <https://doi.org/10.1371/journal.pone.0149985> (2016).
- Krishnamurthy, P. K. & Francis, R. A. A critical review on the utility of DNA barcoding in biodiversity conservation. *Biodiv. Conserv.* **21**, 1901–1919, <https://doi.org/10.1007/s10531-012-0306-2> (2012).
- DeSalle, R. & Amato, G. The expansion of conservation genetics. *Nat. Rev. Genet.* **5**, 702–712, <https://doi.org/10.7312/amat12832-006> (2004).
- Campos, J. C. & Brito, J. C. Mapping underrepresented land cover heterogeneity in arid regions: the Sahara-Sahel example. *ISPRS J. Photogramm.* **146**, 211–220, <https://doi.org/10.1016/j.isprsjprs.2018.09.012> (2018).
- Brito, J. C. *et al.* Armed conflicts and wildlife decline: Challenges and recommendations for effective conservation policy in the Sahara-Sahel. *Conserv. Lett.* **11**(5), e12446, <https://doi.org/10.1111/conl.12446> (2018).
- Weiss, D. J. *et al.* A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336, <https://doi.org/10.1038/nature25181> (2018).
- Geniez, P., Mateo, J. A. & Bons, J. A checklist of the amphibians and reptiles of Western Sahara (Amphibia, Reptilia). *Herpetozoa* **133**, 149–63 (2000).
- Geniez, P., Mateo, J. A., Geniez, M. & Pether, J. *The Amphibians and Reptiles of the Western Sahara. An Atlas and Field Guide*. Chimaira Editions (2004). Available at [https://doi.org/10.1643/0045-8511\(2007\)2007\[772:TAAROT\]2.0.CO;2](https://doi.org/10.1643/0045-8511(2007)2007[772:TAAROT]2.0.CO;2)
- Trape, J.-F. & Mané, Y. *Guide des Serpents d'Afrique Occidentale: Savane et Désert*. IRD éditions (2006).
- Trape, J.-F., Trape, S. & Chirio, L. *Lézards, Crocodiles et Tortues d'Afrique Occidentale et du Sahara*. IRD éditions (2012).
- Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci. USA* **110**, 15758–15763, <https://doi.org/10.1073/pnas.1314445110> (2013).
- Nagy, Z. T., Sonet, G., Glaw, F. & Vences, M. First large-scale DNA barcoding assessment of reptiles in the biodiversity hotspot of Madagascar, based on newly designed COI primers. *PLoS ONE* **7**, e34506, <https://doi.org/10.1371/journal.pone.0034506> (2012).

35. Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **3**, 294–299 (1994).
36. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids. Res.* **32**, 1792–1797, <https://doi.org/10.1093/nar/gkh340> (2004).
37. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120, <https://doi.org/10.1007/BF01731581> (1980).
38. Brown, S. D. *et al.* Spider: An R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. *Mol. Ecol. Res.* **12**, 562–565, <https://doi.org/10.1111/j.1755-0998.2011.03108.x> (2012).
39. Meier, R., Shiyang, K., Vaidya, G. & Ng, P. K. L. DNA barcoding and taxonomy in diptera: a tale of high intraspecific variability and low identification success. *Syst. Biol.* **55**, 715–728, <https://doi.org/10.1080/10635150600969864> (2006).
40. Pizzigalli, C. *et al.* Phylogeographic diversification of the *Mesalina olivieri* species complex (Squamata: Lacertidae) with the description of a new species and a new subspecies endemic from North West Africa. *J. Zool. Syst. Evol. Res.* **59**, 2321–2349, <https://doi.org/10.1111/jzs.12516> (2021).
41. Mediannikov, O., Trape, S. & Trape, J.-F. A molecular study of the genus *Agama* (Squamata: Agamidae) in West Africa, with description of two new species and a review of the taxonomy, geographic distribution, and ecology of currently recognized species. *Russ. J. Herpetol.* **19**, 115–142, <https://doi.org/10.30906/1026-2296-2012-19-2-115-142> (2012).
42. Wagner, P., Wilms, T. M., Bauer, A. & Böhme, W. Studies on African *Agama*. V. On the origin of *Lacerta agama* Linnaeus, 1758 (Squamata: Agamidae). *Bonn. zool. Beitr.* **56**, 215–223 (2009).

Acknowledgements

Field surveys and sample collection made under authorization from the Ministry of Environment and Durable Development of Mauritania (permits 2012-827/MDAPMCEDD/SG, 2015-162/Dir/PND, 2015-100/Dir/PND, 2016-166/D/PND, 2017-063/MEDD/PND/D) and the High Commission for Water and Forests and the Fight Against Desertification of Morocco (permits 278/2012/ and 20/2013/HCEFLCD/DLCPN/DPRN/CFF). Funding provided by National Geographic Society (CRE-7629-04, CRE-8412-08, GEFNE-53-12), Mohammed bin Zayed Species Conservation Fund (11052709, 11052707, 13257467), Rufford Foundation (SG-15399-1, SG-17893-1), Fundação para a Ciência e Tecnologia (FCT: PTDC/BIA-BEC/099934/2008, PTDC/BIA-BIC/2903/2012), FEDER through COMPETE-Operational Programme for Competitiveness Factors (FCOMP-01-0124-FEDER-008917, -028276), and by AGRIGEN-NORTE-01-0145-FEDER-000007, supported by Norte Portugal Regional Operational Programme (NORTE2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF). Work supported by National Funds through FCT in the scope of the project UIDB/50027/2020. Research conducted in the scope of the LIA “Biodiversity and Evolution”. Individual support given by FCT (contracts CEECINST/00014/2018/CP1512/CT0001, DL57/2016/CP1440/CT0010, CEECIND/00937/2018 and IF/01425/2014, and grant SFRH/BD/140348/2018) and by the Spanish Government (Ref. RYC-2019-026959-1/AEI/10.13039/501100011033). Acknowledgements extended to A.S. Sow, D.V. Gonçalves, J.C. Campos, L. García-Cardenete, N. Sillero, P. Tarroso, S. Fahd and X. Santos for sampling support, I. Ineich (MNHN Paris) for additional samples, and D. Castro for laboratory support. Logistic support for fieldwork was given by M.Y.O.M. Lemine (PN Banc d’Arguin), D.O.S.O. Daf, Z.E.A. Sidatt and M. Aveloitt (PN Diawling), S.M.O. Lehlou (Ministère de l’Environnement et du Développement Durable of Mauritania), D. Hamidou (University of Nouakchott), and A. Araújo (MAVA Foundation).

Author contributions

G.V.-A. and J.C.B. designed the study; J.C.B., F.M.-F., G.V.-A., P.G., P.A.-C. and J.P. collected samples; J.C.B. and G.V.-A. contributed reagents; M.H. and A.L. produced barcodes; P.A.-C. and P.G. identified from the photos all specimens before the barcoding; all authors organized data and cleaned the database; G.V.-A. analyzed data; G.V.-A. drafted the manuscript with contributions from all other authors.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to G.V.-A. or J.C.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022